Virus Genes (2010) 40:111–118
DOI 10.1007/s11262-009-0409-3

Genetic diversity and evidence for recent modular recombination in Hawaiian *Citrus tristeza virus*

Michael J. Melzer · Wayne B. Borth · Diane M. Sether · Stephen Ferreira · Dennis Gonsalves · John S. Hu

Received: 6 August 2009 / Accepted: 26 September 2009 / Published online: 16 October 2009

Abstract The Hawaiian Islands are home to a widespread and diverse population of *Citrus tristeza virus* (CTV), an economically important pathogen of citrus. In this study, we quantified the genetic diversity of two CTV genes and determined the complete genomic sequence for two strains of Hawaiian CTV. The nucleotide diversity was estimated to be 0.0565 ± 0.0022 for the coat protein (*CP*) gene ($n = 137$) and 0.0822 ± 0.0033 for the *p23* gene ($n = 30$). The genome size and organization of CTV strains HA18-9 and HA16-5 were similar to other fully sequenced strains of CTV. The 3′-terminal halves of their genomes were nearly identical (98.5% nucleotide identity), whereas the 5′-terminal halves were more distantly related (72.3% nucleotide identity), suggesting a possible recombination event. Closer examination of strain HA16-5 indicated that it arose through recent recombination between the movement module of an HA18-9 genotype, and the replication module of an undescribed CTV genotype.

Keywords dsRNA cloning · Viral genetic diversity · RNA virus recombination · *Closteroviridae*

Electronic supplementary material The online version of this article (doi:10.1007/s11262-009-0409-3) contains supplementary material, which is available to authorized users.

M. J. Melzer · W. B. Borth · D. M. Sether · S. Ferreira · J. S. Hu (✉)
Department of Plant and Environmental Protection Sciences, University of Hawai'i, Honolulu, HI, USA
e-mail: johnhu@hawaii.edu

D. Gonsalves
USDA Pacific Basin Agricultural Research Center, Hilo, HI, USA

Introduction

Citrus tristeza virus (CTV; family *Closteroviridae*, genus *Closterovirus*), an important viral pathogen of citrus worldwide, possesses a ~20 kb single-stranded RNA genome that encodes 12 open reading frames (ORFs) [1, 2]. The two 5′-terminal ORFs encompass roughly half of the genome and primarily encode replication-associated products, whereas the remaining 10 3′-terminal ORFs encode proteins involved with virus movement [3]. The complete nucleotide sequence has been determined for several CTV strains. Phylogenetic analyses using these data have identified three major clades of CTV, which are typified by strains T36, VT, and T30 [4]. Within these clades, nucleotide sequence variation is generally within 5% and is distributed relatively evenly throughout the genome [4, 5]. Between clades, the 3′-terminal half of the genome is well conserved, whereas the 5′-terminal half can exhibit a great deal of sequence divergence [6, 7]. For example, the nucleotide sequence identity between strains T36 and VT in the 5′-untranslated region (5′-UTR) can be as low as 41% [5]. It has thus been suggested that strain T36 might be the descendent of a recombinant strain composed of the 3′-terminal region of a VT- or T30-like genome, and the 5′-terminal region of an undescribed strain of CTV. The transition zone between the 5′- and 3′-terminal regions of T36 is gradual over hundreds of nucleotides, indicating the recombination event did not occur recently [7].

CTV and its most efficient vector, the brown citrus aphid (*Toxoptera citricida* Kirk.), have been present in the Hawaiian Islands for over 50 years [8]. As a result, Hawaii has a diverse CTV population, which includes most of the previously identified genotypes [9]. Many trees also appear to be infected with multiple and

possibly unreported genotypes [9]. In this study, we assess the genetic diversity of CTV in Hawaii using the coat protein (*CP*) and *p23* genes. We also fully sequence the genomes of two CTV strains with non-standard RT-PCR genotype profiles [6, 9, 10] using an adaptation of random PCR (rPCR) [11] and degenerate oligonucleotide primed-PCR (DOP-PCR) [12].

Materials and methods

Assessing the genetic diversity of Hawaiian CTV

The *CP* and *p23* gene sequence data were used to assess the genetic diversity of the Hawaiian CTV population. The CP sequence data were obtained by performing reverse-transcriptase (RT)-PCR on CTV-positive samples that were pooled by island [9]. The primer sequences of T36cp(+) and T36cp(−) used to amplify the *CP* gene [6] were trimmed from this dataset in subsequent analyses. Much of the CP sequence data were previously collected [9], however, additional CP sequence data were obtained in this study for the islands of Kauai (8 clones), Oahu (6 clones), Maui (12 clones), and Hawaii Island (9 clones) using the same protocol. A total of 137 *CP* gene sequences were examined in this study (Table 1).

Sequence data for p23 were obtained in a similar manner using the degenerate primer 389 (5′-AAGTCTGYRA GTTACRATGG-3′) and primer 390 (5′-TTATTCCGTC CACTTCAATC-3′). The underlined sequences for 389 and 390 include the start and partial stop codons of the *p23* gene and were included in analyses of the *p23* gene. Primer sequences not underlined were trimmed before analyses. A total of 30 clones were examined for the islands of Kauai, Oahu, Maui, and Hawaii Island.

Estimates of genetic diversity were generated using DnaSP 4.50.3 [13]. *CP* gene nucleotide sequences were converted to amino acid sequences using the "Translate" function of SeqWeb v. 2 (GCG) to eliminate silent polymorphisms. The neighbor-joining algorithm in the "GrowTree" function of SeqWeb v.2 was used to infer the phylogenetic relationships of these amino acid sequences.

Table 1 Jukes and Cantor nucleotide diversity (π_{JC}) and standard deviation (σ) of *Citrus tristeza virus* in the Hawaiian Islands

Island	CP		p23	
	n	π_{JC} (σ)	*n*	π_{JC} (σ)
Kauai	29	0.0531 (0.0048)	7	0.0739 (0.0120)
Oahu	31	0.0657 (0.0023)	8	0.0766 (0.0107)
Molokai	10	0.0496 (0.0069)	n/a	n/a
Maui	33	0.0199 (0.0067)	7	0.0753 (0.0098)
Hawaii Island	33	0.0483 (0.0024)	8	0.0879 (0.0075)
Total	137	0.0565 (0.0022)	30	0.0822 (0.0033)

Double-stranded RNA source and isolation

In 2001, CTV samples HA18-9 and HA16-5 were collected from Tahitian lime (*Citrus latifolia*) trees located at site HA18 in Waimea and site HA16 in Kailua-Kona, Hawaii Island, respectively. Approximately 4 g and 100 mg of bark tissue were available for dsRNA isolation from samples HA18-9 and HA16-5, respectively, which was insufficient for cloning and sequencing the infecting strain(s) with approaches commonly used for CTV. Additionally, since these samples were collected more than 8 years ago, the possibility that the CTV genotype profiles in these trees has changed since the initial samples were taken precluded the collection of additional tissue from these trees. Instead, rPCR [11] and DOP-PCR [12] approaches were adapted for cloning and sequencing the CTV present in these samples.

Tissue was ground in liquid nitrogen and total nucleic acids (TNAs) were extracted as described [14], except that diethyldithiocarbamic acid was omitted from the extraction buffer. The TNAs were resuspended in 17 µl of H_2O and 2 µl of 10× RQ1 DNase buffer (Promega), then digested with 1 U of RQ1 DNase at 37°C for 60 min. Following digestion, 80 µl of STE (1 M NaCl, 200 mM Tris–HCL, and 100 mM EDTA, pH 8.0), 20 µl of 95% ethanol, and 10–20 mg of CF-11 cellulose (Whatman) were added, and the samples were gently shaken for 1–2 h. The sample was then applied to a centrifuge column containing a GF/C filter (Whatman), centrifuged at 1,000 × *g* for 30 s and the eluate was discarded. The column was washed with three 300 µl aliquots of STE containing ethanol (16.5% v/v) and the eluate was discarded after each wash. STE (100 µl) was added to the column, left to stand for 5 min, and centrifuged. The eluate was adjusted to 16.5% ethanol and 10–20 mg of CF-ll cellulose were added and gently shaken for 1–2 h. The sample was passed through a new GF/C filter centrifuge column, washed as described above, and the dsRNA eluted with 100 µl of STE. This dsRNA eluate was concentrated with a YM-30 or YM-50 centrifuge column (Millipore) and used as template for cDNA synthesis.

cDNA synthesis and molecular cloning strategy

Double-stranded RNAs were heat-denatured at 95°C for 10 min with 10 pmol of the rPCR primer Universal-dN_6 (5′-GCCGGAGCTCTGCAGAATTCNNNNNN-3′) [11] and quickly chilled on ice. A 20 µl reaction was set up containing the denatured dsRNA and Universal-dN6, 50 mM Tris–HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 0.2 mM of each dNTP, and 200 U of SuperScript II

reverse-transcriptase (Invitrogen). The reaction was incubated at 25°C for 10 min, and then at 50°C for 50 min. Approximately 0.75 U of RNase H (Promega) was added to the reaction, incubated at 37°C for 30 min, and then concentrated with either a YM-30 or YM-50 column. The column eluate was used as primer/template in a 20 µl hotstart overlap-extension PCR reaction using 2X ImmoMix (BioLine), which consisted of: 7 min at 95°C; 10 cycles of 60 s at 95°C, 60 s at 55°C, and 60 s at 72°C; and 7 min at 72°C. The overlap-extension PCR reaction (1 µl) was used as template in a 20 µl single-primer PCR reaction containing 20 pmol of the rPCR primer Universal (5′-GCCGGAGCTG TGCAGAATTC-3′) [11] and either Easy-A High-Fidelity PCR Cloning Enzyme (Stratagene) or 2X ImmoMix. The cycling parameters were: 7 min at 95°C; 25–35 cycles of: 60 s at 95°C, 60 s at 58°C, and 60 s at 72°C; and 7 min at 72°C. PCR products were either size-fractionated by excising a desired product size following electrophoresis in a 1% (w/v) agarose gel or directly ligated into the "TA" cloning vector pGEM-T Easy (Promega). Recombinant vectors were sequenced at the University of Hawaii's Advanced Studies of Genomics, Proteomics, and Bioinformatics laboratory.

Genome assembly and phylogenetic analysis

Contiguous sequences were assembled using the CAP3 Sequence Assembly program [15]. When gaps in sequence data occurred, degenerate primers based on other CTV strains or primers flanking the gaps were designed and added to a PCR reaction containing 0.5 µl of the overlap-extension PCR reaction as template. PCR products were generated, cloned, and sequenced as above, with 3–5 clones sequenced for each amplicon. The small amount of citrus tissue for sample HA16-5 did not yield enough dsRNA template for characterization of the terminal sequences of the CTV present using either polyadenylation or RACE procedures. For the genomes of CTV sequenced to date, the 3′-termini are well conserved, and the 5′-termini have only a few sequence variations [16]. Therefore, to obtain the 3′-terminal sequences by PCR, primer 702 (5′-ctcaccaacag tagcttATGTTGGCCCCCCATAGG-3′; lower-case bases are not virus-specific), which anneals to the 3′-terminus of most CTV was coupled with a dsRNA-specific primer designed from a clone that mapped close to the 3′-terminus. A similar strategy was employed to obtain the 5′-terminus, in which primers "T30-5′(+)," "T36-5′(+)," or "VT-5′(+)" [6] representing different variations of the CTV 5′-terminus, were individually coupled with a dsRNA-specific primer designed from a clone close to the 5′-terminus. These primer combinations were added to a PCR reaction containing 0.5 µl of the overlap-extension PCR reaction as template, and amplification products were cloned and sequenced as described above. At least five clones were sequenced for each terminus.

Sequences were aligned and neighbor-joined phylogenetic trees were generated using ClustalX v1.8 [17]. Phylogenetic trees were also generated using Bayesian analyses with MRBAYES v3.1.1 [18, 19]. Other CTV sequences used in the phylogenetic analyses were obtained from GenBank: B165 (accession EU076703), a strain from Mexico (DQ272579), NUagA (AB046398), NZ-B18 (FJ525436), NZRB-G90 (FJ525432), NZRB-M12 (FJ525431), NZRB-M17 (FJ525435), NZRB-TH28 (FJ525433), NZRB-TH30 (FJ525434), Qaha (AY340974), SY568 (AF001623), T30 (AF260651), T36 (U16304), T318A (DQ151548), T385 (Y18420), and VT (U56902).

Results

Genetic diversity in Hawaiian CTV

The *CP* and *p23* genes of CTV were used to assess its genetic diversity in Hawaii. Of the 626 nucleotide positions of the *CP* gene analyzed, 199 (32%) were polymorphic. It appears that the island of Maui has the least genetically diverse CTV population, whereas Oahu's population was most diverse (Table 1). This assessment was supported by phylogenetic analysis of the translation of these sequences. Sequences from the Oahu population were distributed throughout the tree, whereas sequences from Maui were more localized (Figure S1). Of the 630 nucleotide positions of the *p23* gene analyzed, 183 (29%) were polymorphic. The nucleotide diversity values of the *p23* gene ranged from 0.0739 to 0.0879 for the different islands (Table 1).

HA18-9

The sequence of 66 unique clones from the HA18-9 sample library indicated that only a single CTV strain was present, which was designated as strain HA18-9 (GenBank Accession # GQ454869). These clones provided 2–8X coverage of 69% of the genome. The sequence gaps representing the remaining 31% of the genome were obtained by PCR. The genome of strain HA18-9 was found to be 19,245 bp in length, and was very similar to the organization of other CTV genomes. The genome of strain HA18-9 was most similar to those of the NZRB group of CTV (Table 2). Comparisons across the entire genomes revealed that the nucleotide sequence identities between strain HA18-9 and the five NZRB strains ranged from 94.3 to 95.5%. Results similar to the nucleotide comparisons were obtained when the proteomes (combined amino acid sequence of the ORFs) were compared (Table 2).

Table 2 Nucleotide (nt; upper right) and amino acid (aa; lower left) percent identity comparisons between fully sequenced strains of *Citrus tristeza virus* (CTV). Nucleotide and amino acid comparisons are based upon the entire genome and proteome, respectively. HA16-5 is a putatively recombinant strain of CTV with the recombination site at the 3′-terminus of the RdRp ORF. Therefore, strain HA16-5 was compared using the 5′- and 3′-terminal regions of the genome on either side of the recombination site, as well as the full-length (FL) sequences

aa	nt																		
	HA16-5			HA18-9	NZRB M17	NZRB TH30	NZRB G90	NZRB TH28	NZRB M12	T36	Qaha	Mexico[a]	NZ B18	B165	T385	T30	T318A	NUagA	VT	SY568
	5′	3′	FL																	
HA16-5																				
5′				72.3	72.0	72.0	72.5	72.6	72.5	71.9	71.2	71.3	82.0	81.8	78.2	78.1	78.4	77.5	77.3	77.7
3′				98.4	92.3	92.5	92.2	92.5	92.5	89.8	89.4	89.7	90.5	90.2	90.1	90.1	90.7	90.8	89.9	90.4
FL				84.1	80.8	80.9	81.0	81.2	81.2	80.2	79.7	79.8	85.6	85.5	83.5	83.4	83.9	83.7	83.0	84.0
HA18-9	75.5	98.5	85.1		94.3	94.3	95.3	95.5	95.5	90.1	89.5	89.2	80.6	80.7	81.3	81.3	81.3	81.0	80.3	82.1
NZRB M17	75.4	93.5	83.0	94.5		98.4	95.4	95.7	95.7	91.5	90.3	89.8	80.6	80.6	81.8	81.7	81.2	80.8	81.1	81.6
NZRB TH30	75.4	93.8	83.0	94.5	98.3		95.5	96.9	96.5	91.5	90.3	89.9	80.7	80.7	81.9	81.8	81.3	80.8	81.2	81.6
NZRB G90	75.6	93.1	82.9	95.4	95.0	95.0		96.6	96.7	90.4	89.3	88.8	80.6	80.8	82.0	82.1	81.5	80.8	81.2	81.7
NZRB TH28	75.5	94.0	83.2	95.8	95.6	96.5	96.4		99.0	90.4	89.3	88.9	80.7	80.8	82.1	82.1	81.4	80.8	81.3	81.6
NZRB M12	75.3	93.9	83.0	95.7	95.5	96.0	96.2	98.7		90.4	89.3	88.9	80.5	80.7	82.1	82.1	81.4	80.7	81.2	81.6
T36	70.7	92.2	79.6	88.9	90.1	90.0	89.0	89.3	89.1		99.1	98.4	80.1	80.3	81.3	81.3	80.6	80.2	79.5	82.0
Qaha	70.0	91.1	78.7	87.9	88.8	88.8	87.8	88.1	88.0	98.3		97.3	79.2	79.8	80.8	80.8	80.2	79.8	79.0	81.6
Mexico[a]	na	na	na	na	na	na	na	na	na	na	na		79.0	79.6	80.8	80.8	80.2	79.8	79.1	81.6
NZ B18	84.1	92.8	87.7	82.9	83.1	83.1	83.2	83.1	83.0	80.2	79.0	na		98.1	86.5	86.5	91.0	90.2	91.2	88.1
B165	83.3	92.5	87.1	82.7	83.2	83.3	83.4	83.4	83.2	80.3	79.2	na	97.9		87.0	87.0	91.6	91.0	90.8	89.7
T385	81.2	92.8	86.0	83.8	84.2	84.3	84.1	84.3	84.0	81.0	80.1	na	89.6	89.7		99.3	90.4	89.8	89.3	94.0
T30	81.1	93.1	86.1	83.7	84.3	84.4	84.1	84.3	84.1	81.2	80.0	na	89.9	90.1	98.9		90.4	89.8	89.3	94.0
T318A	81.8	93.8	86.7	84.2	84.0	84.1	84.4	84.3	84.1	81.1	80.3	na	92.9	92.8	93.1	93.0		95.9	95.2	95.9
NUagA	79.8	93.7	85.6	83.1	83.2	83.2	83.4	83.3	83.1	80.3	79.5	na	91.5	91.4	91.1	91.1	96.6		94.3	94.9
VT	80.8	93.4	86.1	83.6	83.8	83.9	84.1	84.2	84.0	80.8	79.6	na	93.3	93.2	92.7	93.0	96.8	94.8		93.3
SY568	80.8	92.9	85.8	83.7	84.2	84.2	84.4	84.4	84.2	81.4	80.2	na	90.8	90.9	94.2	94.6	96.0	93.8	95.0	

[a] The accuracy of the predicted amino acid sequence is suspect

HA16-5

The sequence data of 95 unique clones from sample HA16-5 indicated that three CTV strains were present in this sample. One set of clones had high (≥ 96%) nucleotide identity with members of the VT clade (e.g., VT, SY568, and NUagA). A second set of clones had very high (> 98%) nucleotide identity with strain HA18-9. A third set of clones had low (70–87%) nucleotide identity with other fully sequenced CTV strains. The first and second sets of clones were evenly distributed throughout the CTV genome, whereas the third set of clones mapped only to the 5′-half of the genome. Clone 295, which spanned from the RdRp ORF to the p33 ORF, indicated that a recombinant strain of CTV was present in sample HA16-5 (Fig. 1). This putatively recombinant strain was designated HA16-5 (GenBank Accession # GQ454870). The two parental sequence donors of strain HA16-5 appear to be strain HA18-9 for the 3′-half of the genome, and an uncharacterized CTV strain for the 5′-half. The putative point of recombination maps to near the 3′-terminus of the RdRp ORF, and produces a truncated RdRp which, when compared to other CTV strains, lacks the five C-terminal residues (Fig. 1). The 10 C-terminal codons of the RdRp ORF are repeated immediately downstream of the RdRp in a −1 frameshift. Since these sets of codons are not identical, it appears that one set may have originated from the donor molecule, and the other from the acceptor molecule.

To confirm that strain HA16-5 was indeed recombinant, and not an artifact of the cloning procedure, other trees from site HA16 were analyzed for the presence of the recombinant strain. TNAs from four trees at this site (HA16-1, HA16-4, HA16-7, and HA16-10) were used as template in RT-PCR analyses using specific primers that flanked the putative recombination site. Faint amplicons of the expected size were generated from two of the four samples. The amplicon from sample HA16-1, however, was the most prominent (data not shown) and was selected for cloning and sequencing. The two sequenced clones were >99% identical to the sequence of strain HA16-5,

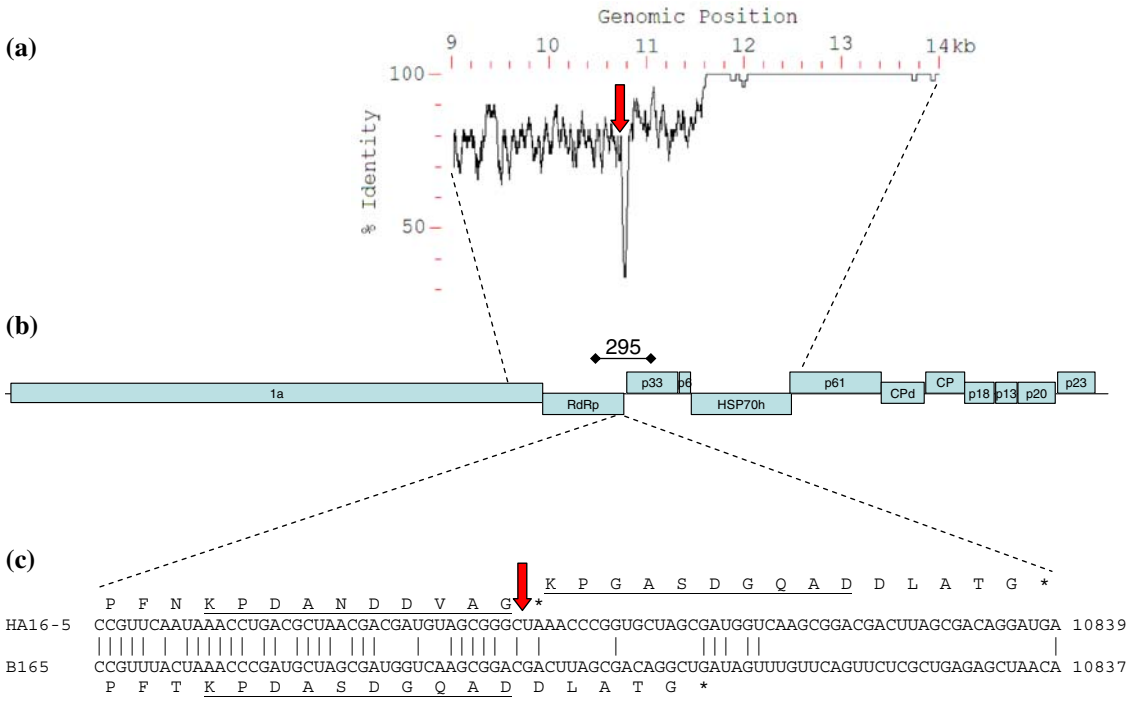

Fig. 1 **a** Sequence comparison plot between the *Citrus tristeza virus* (CTV) strains HA16-5 and HA18-9 for a 5 kb region encompassing the recombination site. The plot was generated by PlotSimilarity function of SeqWeb version 2 (GCG) using a 50 bp comparison window. **b** Genome organization of CTV strain HA16-5. *Boxes* represent open reading frames (ORFs), with the box offset indicating their respective reading frame (e.g., p33, 1a, and RdRp are in the +1, +2, and +3 reading frame, respectively). The position of clone 295, which spans the putative recombination site is indicated. *1a* ORF1a, which contains protease, methyltransferase, and helicase domains; *RdRp* RNA dependent RNA polymerase; *HSP70h* heat shock protein 70 homolog; *CPd* coat protein duplicate; *CP* coat protein. **c** Sequence comparison between strains HA16-5 and B165 at the C-terminal end of the RdRp. Amino acid translations immediately above or below the nucleotide sequence represent the C-terminus of the RdRp. The amino acid translation two levels above the strain HA16-5 nucleotide sequence represents the C-terminus of the donor molecule's RdRp, but it is out of frame and presumably not expressed. The underscored amino acid residues depict homologous sequences that are derived from the donor and acceptor molecules that may have recombined to form strain HA16-5. The *arrows* indicate the putative recombination site in strain HA16-5 and the *dashed lines* map the sequence comparisons to the CTV genome for **a** and **c**

indicating that this recombinant sequence was not an artifact of cloning, since it was present in at least one other tree from the site.

The sequence of strain HA16-5 at the 5′-end of the recombination site had low homology to other fully sequenced CTV strains; 71.2–81.8% and 70.0–83.3% identities for nucleotide and amino acid sequences, respectively (Table 2). Conversely, sequences at the 3′ end of the recombination site were very similar to those of other fully sequenced strains of CTV. Aside from strain HA18-9, which is >98% identical to HA16-5 in this region for both nucleotide and amino acid sequences, identities were between 89.4–92.5% and 91.1–94.0% for nucleotide and amino acid sequences, respectively (Table 2). Surprisingly, the full-length sequence of HA16-5 was most similar to strain NZ-B18, despite the high homology that the 3′-end of strain HA16-5 has with strain HA18-9 (Table 2).

Phylogenetic analyses

The complete genomes and proteomes of strains HA18-9, HA16-5, and other fully sequenced CTV strains were analyzed to elucidate their phylogenetic relationships. The CTV strain from Mexico was omitted from amino acid analyses due to aberrations in this data, which appear to be due to sequencing and/or translation errors. Additional analyses were performed using sequence data upstream and downstream from the putative recombination site of strain HA16-5. Analyses using the nucleotide and amino acid sequences, regardless of what region of the genome or proteome was being analyzed, produced nearly identical results, all with high branch support (Fig. 2; data not shown).

When entire genomic or proteomic sequences or only the sequences upstream from the putative point of recombination were analyzed, strain HA18-9 was found to group with the NZRB group of CTV, whereas strain HA16-5 was placed on an intermediate branch between the two main groups of CTV (Fig. 2). When sequences downstream from the putative point of recombination were analyzed, five distinct clades were created: (1) HA18-9, HA16-5; (2) T36, Qaha, Mexico; (3) T30, T385, SY568; (4) VT, NUagA, T318A, B165, NZ-B18; and (5) the NZRB group (Fig. 2).

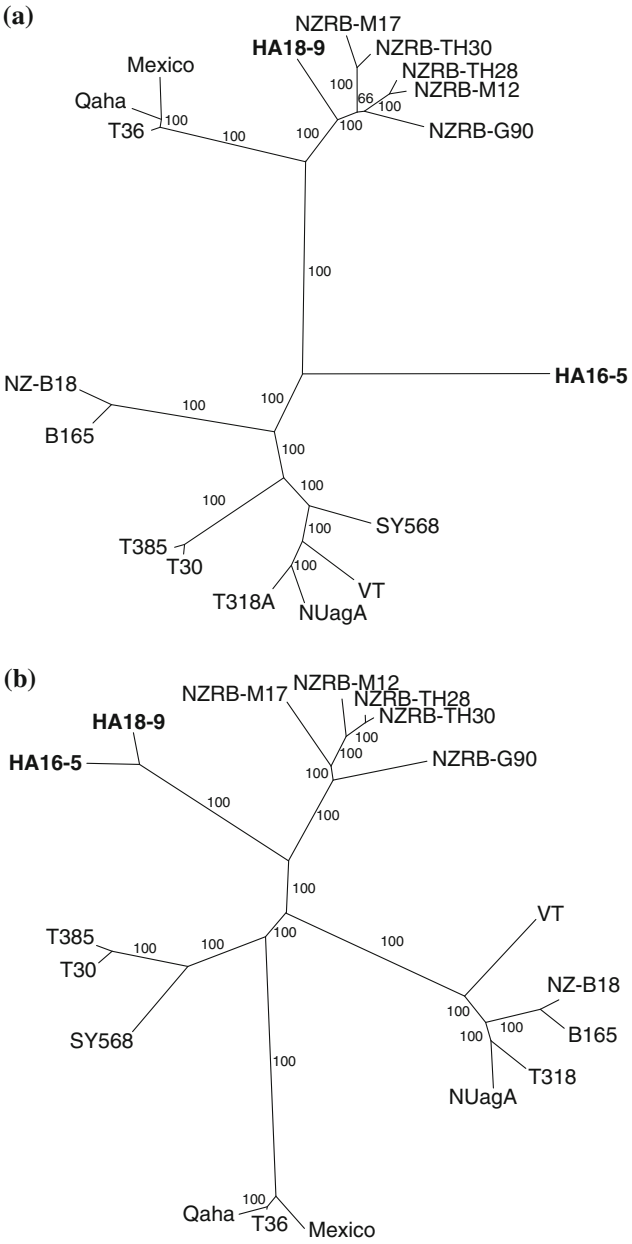

Fig. 2 Phylogenetic analyses of the completely sequenced strains of *Citrus tristeza virus* (CTV). Unrooted neighbor-joining trees are based on the nucleotide sequences of the full-length (**a**) and the 3′-terminal half (roughly nts 10,800-3′-terminus) (**b**) of the genomes. *Numbers* are bootstrap percentages after 1,000 replications, and represent branch support of the trees. The tree generated using the 5′-terminal half of the genomes was similar to the tree depicted in **a** (data not shown). Amino acid translations of the ORFs encoded by these sequences produced essentially identical trees (data not shown)

Discussion

HA18-9 and HA16-5 represent the first strains of CTV from Hawaii to be completely sequenced, and add to the known diversity of this economically important virus. Due to the limited amount of plant tissue available from sample HA16-5, a small-scale dsRNA isolation procedure, coupled with a dsRNA cloning procedure requiring low amounts of dsRNA template was developed. Less than 100 mg of CTV-infected tissue was all that was required for genomic sequencing using this approach. Overall, this approach was very efficient, and proved to be a viable method of genomic sequencing for CTV or other dsRNAs when the amount of

material is limited. Rott and Jelkmann [20] and Benthack et al. [21] also used a modified DOP-PCR procedure to obtain sequences of uncharacterized dsRNAs. In their approach, primers that were degenerate at internal positions were used and may have restricted the number of potential annealing sites on the dsRNA. These primers produced discrete amplicons, some of which were specific to the dsRNA template, but many others were not [20, 21]. In these studies, DOP-PCR was used to obtain initial sequence data and other cloning methods were used to obtain the rest. It is likely that the location of degeneracy in an effective DOP-PCR primer is dependent upon the size of the template molecule being targeted. For large molecules, such as prokaryotic or eukaryotic DNA genomes, the degeneracy can shift toward the internal positions of the primer since the fixed positions at the 3′-terminus of the primer would have more annealing opportunities in a large genome. Conversely, for smaller molecules, such as viral dsRNAs, the degeneracy should shift toward the 3′-terminus of the primer to create more annealing opportunities.

Strain HA18-9 was most similar to members of the NZRB group and together they represent a new genotype of CTV. This genotype appears to be common in Hawaii [9], and was found in both samples HA18-9 and HA16-5. The NZRB group of CTV is of interest because they can infect *Poncirus trifoliata*, a citrus relative once thought to be immune to CTV [22]. It is unknown if strain HA18-9 is capable of infecting *P. trifoliata*, since this rootstock is not commonly used in Hawaii.

Sample HA16-5 contained three strains of CTV. One of them, designated strain HA16-5, represents an intriguing strain of CTV. Not only does it represent a new CTV genotype, but this strain also appears to have arisen through recombination between the 3′-end of a CTV strain of the HA18-9/NZRB genotype and the 5′-end of an undescribed CTV strain. This scenario has also been proposed for the origin of strain T36 [7], although a major difference exists between strains T36 and HA16-5. The transition between the 5′- and 3′- halves of the T36 genome is gradual, which makes it difficult to identify the exact point of recombination, suggesting that this event did not occur recently [7]. In strain HA16-5, however, the putative recombination site is easily identified since it appears that both the donor and acceptor molecules contributed a 31-nt stretch of sequence that resulted in a heterogenous tandem repeat that flanks the recombination site. This sequence redundancy suggests that the recombination event occurred relatively recently. Although it appears that an HA18-9/NZRB genotype contributed the 3′-terminal half of the genome of strain HA16-5, the initial ~800 bp of sequence downstream from the putative recombination site, which represents most of the p33 ORF, has low similarity to strain HA18-9 (Fig. 1). This ~800 bp sequence might indicate another recombination event, similar to those described in other CTV strains [23–25].

Nagy [26] has proposed four classes of virus recombination. Class 2 and 3 recombinants require the presence of *cis*-acting control elements either without or with the assistance of base-pairing, respectively. In CTV, these control elements are composed of essential primary and secondary structures [27]. The putative recombination site in HA16-5 is located two nucleotide positions upstream from the *cis*-acting control element secondary structure that is required for the genesis of the 5′-terminal LaMT (Large Molecular-weight Tristeza) subgenomic (sg)RNA [28], and may also be close to that of the 3′-terminal p33 sgRNA [27]. It is, therefore, possible that the 5′-terminal LaMT sgRNAs of an unknown strain of CTV, or the p33 sgRNA of an HA18-9 genotype, were donors or acceptors in a class 2 or 3 recombination event.

The 5′-terminal half of the genome of strain HA16-5 has relatively low sequence homology to any characterized strain of CTV. This region of the genome encodes the replication-associated proteins, while the 3′-terminal half of the genome encodes proteins involved in assembly and movement of the virus in citrus. CTV appears to replicate in the cells of all citrus species, but varies in its ability to move cell-to-cell and long distance [29, 30]. This variability in movement depends on the strain of CTV and the host species it infects. It has been proposed (M. Hilf, personal communication) that if an uncharacterized strain of CTV that is able to replicate and move in one citrus species is acquired and deposited by an aphid vector into another species, it may be able to replicate in but not move out of the initially infected cell. If this cell was also infected with a movement-competent strain of CTV, recombination between the 5′-terminal half of the genome (replication-associated proteins) of the movement-incapable strain with the 3′-terminal half of the genome (movement-associated proteins) of the movement-capable strain, would produce a chimeric CTV strain that would be able to replicate and move in the host. Such a recombinant strain would have a genome structure similar to that of strain HA16-5. In fact, the replication module of any *Closterovirus* capable of replicating in citrus cells could undergo this type of modular recombination.

The sequence diversity of the *CP* and *p23* genes, and the characterization of two strains of CTV, supports previous observations of high genetic diversity in Hawaiian CTV [9]. Both of these sequenced strains represent new genotypes of the virus, one of which appearing to be recombinant. Interestingly, it also appears that one of the strains has donated genetic material to the other. Recombination is an important strategy for the evolution of new viral genotypes. It is likely that recombination has also added to the genetic diversity of the CTV population in Hawaii.

Acknowledgments The research was funded, in part, by grants from USDA-CSREES T-STAR program agreement #2003-34135-13982 and a specific cooperative grant agreement #58-5320-5-785, between the USDA-ARS Pacific Basin Agricultural Research Center and the University of Hawaii at Manoa.

References

1. H.R. Pappu, A.V. Karasev, E.J. Anderson, S.S. Pappu, M.E. Hilf, V. Febres, R.M.J. Eckloff, M. McCaffery, V. Boyko, S. Gowda, V.V. Dolja, E.V. Koonin, D.J. Gumpf, K. Cline, S.M. Garnsey, W.O. Dawson, R.F. Lee, C.L. Niblett, Virology **199**, 35 (1994)
2. A.V. Karasev, V.P. Boyko, S. Gowda, O.V. Nikolaeva, M.E. Hilf, E.V. Koonin, C.L. Niblett, K. Cline, D.J. Gumpf, R.F. Lee, S.M. Garnsey, D.J. Lewandowski, W.O. Dawson, Virology **208**, 511 (1995)
3. V.V. Dolja, J.F. Kreuze, J.P.T. Valkonen, Virus Res. **117**, 38 (2006)
4. P. Moreno, S. Ambrós, M.R. Albiach-Martí, J. Guerri, L. Peña, Mol. Plant Pathol. **9**, 251 (2008)
5. M.R. Albiach-Marti, M. Mawassi, S. Gowda, T. Satyanarayana, M.E. Hilf, S. Shanker, E.C. Almira, M.C. Vives, C. Lopez, J. Guerri, R. Flores, P. Moreno, S.M. Garnsey, W.O. Dawson, J. Virol. **74**, 6856 (2000)
6. M.E. Hilf, A.V. Karasev, M.R. Albiach-Marti, W.O. Dawson, S.M. Garnsey, Phytopathology **89**, 336 (1999)
7. M. Mawassi, E. Mietkiewska, R. Gofman, G. Yang, M. Bar-Joseph, J. Gen. Virol. **77**, 2359 (1996)
8. D.C. Giacometti, W.B. Storey, Cal. Citrograph **37**, 357 (1952)
9. M.J. Melzer, W.B. Borth, F. Zee, M.E. Hilf, S.M. Garnsey, J.S. Hu, in *Proceedings of the 16th Conference of the International Organization of Citrus Virologists*, ed. by M.E. Hilf, N. Duran-Vila, M.A. Rocha-Peña (IOCV, Riverside, CA, 2005), p. 179
10. M.E. Hilf, S.M. Garnsey, in *Proceedings of the 14th Conference of the International Organization of Citrus Virologists*, ed. by J.V. da Graça, R.F. Lee, R.K. Yokomi (IOCV, Riverside, CA, 2000), p. 18
11. P. Froussard, Nucleic Acids Res. **20**, 2900 (1992)
12. H. Telenius, N.P. Carter, C.E. Bebb, M. Nordenskjöld, B.A. Ponder, A. Tunnacliffe, Genomics **13**, 718 (1992)
13. J. Rozas, J.C. Sánchez-DelBarrio, X. Messeguer, R. Rozas, Bioinformatics **19**, 2496 (2003)
14. C.N. Stewart Jr., L.E. Via, BioTechniques **14**, 748 (1993)
15. X. Huang, A. Madan, Genome Res. **9**, 868 (1999)
16. C. López, M.A. Ayllón, J. Navas-Castillo, J. Guerri, P. Moreno, R. Flores, Phytopathology **88**, 685 (1998)
17. J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, Nucleic Acids Res. **24**, 4876 (1997)
18. J.P. Huelsenbeck, F. Ronquist, Bioinformatics **17**, 754 (2001)
19. F. Ronquist, J.P. Huelsenbeck, Bioinformatics **19**, 1572 (2003)
20. M.E. Rott, W. Jelkmann, Eur. J. Plant Pathol. **107**, 411 (2001)
21. W. Benthack, N. Mielke, C. Büttner, H.P. Muhlbach, Arch. Virol. **150**, 37–53 (2004)
22. M.E. Hilf, in *Proceedings of the 16th Conference of the International Organization of Citrus Virologists*, ed. by M.E. Hilf, Duran-Vila, M.A. Rocha-Peña (IOCV, Riverside, CA, 2005), p. 52
23. L. Rubio, M.A. Ayllón, P. Kong, A. Fernandez, M. Polek, J. Guerri, B.W. Falk, J. Virol. **75**, 8054 (2001)
24. M.C. Vives, L. Rubio, C. López, J. Navas-Castillo, M.R. Albiach-Marti, W.O. Dawson, J. Guerri, R. Flores, P. Moreno, J. Gen. Virol. **80**, 811 (2005)
25. Z. Weng, R. Barthelson, S. Gowda, M.E. Hilf, W.O. Dawson, D.W. Galbraith, Z. Xiong, PLoS ONE **2** (2007). doi: 10.1371/journal.pone.0000917
26. P.D. Nagy, in *Plant virus evolution*, ed. by M. Roossinck (Springer, Berlin, 2008), p. 133
27. M.A. Ayllón, S. Gowda, T. Satyanarayana, W.O. Dawson, Virology **322**, 41 (2004)
28. X. Che, D. Piestun, M. Mawassi, G. Yang, T. Satyanarayana, S. Gowda, W.O. Dawson, M. Bar-Joseph, Virology **283**, 374 (2001)
29. M.R. Albiach-Marti, J.W. Grosser, S. Gowda, M. Mawassi, T. Satyanarayana, S.M. Garnsey, W.O. Dawson, Mol. Breed. **14**, 117 (2004)
30. S.Y. Folimonova, A.S. Folimonov, S. Tatineni, W.O. Dawson, J. Virol. **82**, 6546 (2008)